

令和 4 年 5 月 24 日現在

機関番号：32642

研究種目：基盤研究(C) (一般)

研究期間：2018～2021

課題番号：18K11361

研究課題名(和文) 集団・個人の階層的認識に基づく人物照合

研究課題名(英文) Hierarchical Person Identification using Individual and Group Features

研究代表者

杉村 大輔 (Sugimura, Daisuke)

津田塾大学・学芸学部・准教授

研究者番号：10712052

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：観測された人物が「誰か」を識別する、人物照合技術について検討する。従来手法は、人物個人の色情報や形状特徴を用いて照合している。しかしながら、人物個人から抽出できる情報は限定的であり、照性能の向上には限界がある。また、照合対象人数が増大した場合、照合性能低下を引き起こす。本研究課題では、集団・個人の階層的照合手法を検討する。集団の情報を解析することができれば、集団に関する特徴量を利用することができる。さらに、集団を照合できれば、この中から個人を照合すれば良いため、照合対象者の数を限定することができる。このような階層的照合アプローチに基づき、複数カメラ間人物照合、個人認証技術の開発を行う。

研究成果の学術的意義や社会的意義

映像からの人物照合技術の高度化は、安心・安全な社会環境の実現において必要不可欠である。人物個人だけでなく集団の情報を照合に活用する観点は、これまでの従来技術にないものである。また、集団・個人の階層的認識は、照合対象の増大に伴う性能低下という当該分野の本質的な課題の解決を試みる独創的なアプローチである。このように、本研究課題で検討および開発した技術は、当該分野にインパクトを与えるものであると考えられる。

研究成果の概要(英文)：We develop a method for person identification from surveillance videos. Conventional methods used color and shape features of individual person. However, the information that can be extracted from individual tends to be limited; thus, there is a certain limit to the improvement of identification performance. In addition, when the number of people to be identified increases, the identification performance generally deteriorates. In this research, we investigate a hierarchical identification approach for groups and individuals. If we can analyze the information of a group, we can use the features related to the group. Furthermore, if the group can be identified, the number of candidates to be identified can be decreased because it can be considered that groups contain only a few people. Based on this hierarchical identification approach, we develop techniques of a person re-identification and person identification.

研究分野：画像認識

キーワード：知覚情報処理 画像認識

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

安心・安全な社会環境の実現のために、映像を用いた人物行動解析技術の確立が期待されている。特に、観測された人物が「誰か」を識別する、人物照合、および個人認証について幅広く研究されてきている。

従来手法の多くは、人物の色情報や形状、深層学習により得られる特徴量を用いて、個人を認識している。しかしながら、人物個人の映像から抽出できる情報は限定的であり、照合性能の向上には限界があると考えられる。たとえば、広域監視の用途を考えた場合、カメラ映像は低解像度・低画質で撮影されることが多い。これはすなわち、従来方式では個人識別するための特徴量の抽出が難しいことを示唆している。これに加え、照合対象人数が増大した場合、登録された個人データ(ギャラリー)が多様になる。このことから、入力された人物画像を正しく照合することがより困難になる。

2. 研究の目的

本研究課題では、個人を直接照合する従来手法とは異なる、集団・個人の階層的な照合手法を検討する。人物は特定の集団(グループ)(例:同僚、友人等)に属し、一緒に行動することが多いと考えられる。これはすなわち、映像中の集団の情報を解析することができれば、個々の人物に関する特徴量だけでなく、集団に関する特徴量を取得、および利用することができることを示唆している。さらに、集団を認識・照合できれば、照合された集団の中から個人を照合すれば良いため、探索するギャラリーの大きさを限定することができる。これらの利点は、人物個人照合の性能向上に寄与すると考えられる。

そこで本研究課題では、映像からの集団の認識・照合を伴う階層的照合技術を開発する。本研究課題で検討する階層的照合アプローチは、様々な個人認証技術への応用が期待できる。そこで、(1)複数カメラ間人物照合、(2)動作特徴を用いた個人認証技術の開発を行う。

3. 研究の方法

(1) 複数カメラ間人物照合

人物照合:

人物集団の見え特徴を活用した個人照合手法を開発する。人物集団は、同時刻に観測された人物群(同僚や友人等)によって構成されると考えられる。各カメラの観測領域が大きく離れていないことを仮定すると、集団を構成する人物群の関係もまたカメラ間で一貫して観測されることが考えられる。これはすなわち、個人の人物の見え特徴に加え、周囲の複数人物の見え特徴も同時に利用できることを示唆している。本検討では、集団・個人の特徴を活用することで、人物照合性能の向上を実現する。

検討した方式の流れは以下の通りである。最初に、個人の人物画像から見え特徴を取得し、個人の見えに基づく個人間距離を算出する。さらに、集団の見え特徴を用いるために、個人と周囲の人物の関係を表現する集団らしさを推定する。得られる集団らしさを重みとして、Earth Mover's Distanceを用いて集団間距離を計算する。最後に、計算された個人間・集団間距離に基づき、カメラ間人物照合を行う。

人物個人照合性能向上に向けた人物移動軌跡推定:

カメラ映像において頑健な個人照合を実現するためには、人物の移動軌跡の推定性能を向上させる必要がある。そこで、グラフニューラルネットワークを用いた移動軌跡片の対応付けに基づく移動軌跡推定手法を開発する。

本検討では、各人物の特徴をノード、人物間の関係をエッジとした時空間グラフ表現を行う。これにより、歩行者移動軌跡推定という問題を、グラフにおいてエッジが存在するか否かを判別するリンク予測問題として扱うことができる。各ノードが持つ情報の特徴量に変換するグラフニューラルネットワークを用いることで、ノード間の相互作用や関係性を考慮した、より識別性の高い特徴量を得ることができる。このようなグラフニューラルネットワークとリンク予測の枠組みを組み合わせることで、人物特徴および関連性を考慮した高精度な歩行者移動軌跡推定を実現する。

(2) 動作特徴を用いた個人認証

瞬き動作認証:

人物の瞬き動作を用いた個人認証手法について検討する。瞬き動作を用いた個人認証手法についてこれまでいくつかの手法が提案されている。しかしながら、カメラと被験者の距離が遠くなるとその認証性能が低下する。これに加え、認証すべき人物の数(ギ

ギャラリー)が大きくなると、認証性能は低下する。

本検討では、遠距離において最終的な識別を行うのではなく、遠距離・近距離における多段階識別手法を開発する。まず、低解像度で撮影される遠距離での瞬き動作より取得した大局的な特徴を用いて、瞬き動作が大まかに類似したグループ単位での識別を行う。次に、近距離での瞬き動作より取得した詳細な特徴を用いることで、識別されたグループ内において個人認証を行う。遠距離におけるグループ単位での認証を行うことで、近距離での個人認証においてギャラリーのサイズを限定することができる。これにより、最終的な瞬き認証性能の向上を実現する。

顔特徴と読唇処理に基づく個人認証：

個人認証性能のさらなる向上を目指すために、瞬き動作に加え、顔特徴と読唇処理に基づく階層的個人認証手法について検討する。高精度な認証方式の一つとして顔認証がよく知られている。しかしながら、顔情報は静的特徴であるため、例えば顔画像を入力することで、不正に認証が突破されてしまう危険性がある。一方、動的特徴の利用による認証方式は、個人の持つ行動パターンや癖を利用するため、静的特徴を用いた認証の問題を抑制することができる。その一つとして、音声認識技術を考える。これは、事前登録された言葉(セキュリティコード)と発話内容の照合を行うことで認証を行っている。しかしながら、環境的要因(雑音環境下による認識精度の低下)と、身体的要因(病気等による音声特徴の変化)に起因し、認識性能低下の恐れがある。

本検討では、顔特徴と発声が必要ない読唇処理に基づいた階層的生体認証システムを開発する。高精度な顔特徴を用いた認証と、なりすましを防ぐ動的な唇動作の解析による認証を組み合わせることで、より頑健な個人認証を実現する。

4. 研究成果

(1) 複数カメラ間人物照合

人物照合の性能評価：

開発した人物照合手法の有効性を確認するために、公開データセット [1, 2] を用いて評価実験を行った。具体的には、NLPR_MCT_Dataset1 [1] に含まれる Cam1-Cam2 シーケンス、SAVIT_SoftBio_Dataset [2] に含まれる Cam3-Cam8 および Cam5-Cam8 シーケンスを用いた。図 1 に各データセットのシーケンスの例を示す。比較手法は、Hierarchical Gaussian Descriptors [3] とした。これは、人物一名が写った一枚画像を用いたカメラ間人物照合手法である。そのため、人物映像からフレーム選択を行う必要がある。そこで、映像からランダムにフレーム選択をし、照合を行った。評価尺度には累積照合特性を用いた。累積照合特性は、全ての人物に対して照合結果の順位付けを行った際、同一人物が上位 R 人以内に照合された割合を意味する。

表 1 に評価実験結果を示す。いずれの場合においても提案手法が高精度であることが確認できる。

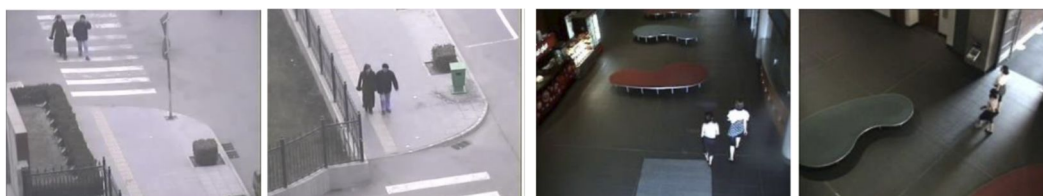


図 1 評価に用いた映像の例：(a) SAVIT_SoftBio_Dataset[1]; (b) NLPR_MCT_Dataset1[2]

表 1 累積照合特性の比較 [%]

Dataset	NLPR_MCT_Dataset1[1]			SAVIT_SoftBio_Dataset[2]					
	Cam1 - Cam2			Cam3 - Cam8			Cam5 - Cam8		
Seq. ID	R=1	R=5	R=10	R=1	R=5	R=10	R=1	R=5	R=10
手法[3]	56.2	87.7	95.9	45.8	72.9	88.1	9.8	32.8	52.5
提案手法	72.6	91.8	95.9	49.2	84.7	93.2	19.7	45.9	63.9

人物移動軌跡推定の性能評価

人物追跡データセット MOT17 [4, 5] を用いて評価実験を行った。人物移動軌跡推定の評価指標として、MOTA および IDF1 を用いた。いずれの評価指標も 1 に近い値であるほど高性能であることを意味する。既存手法である tracklet association tracker [6] と性能比較を行った。

結果、手法 [6] によるものは、MOTA が 0.315、IDF1 が 0.371 であった。一方提案手法によるものは、MOTA が 0.334、IDF1 が 0.377 であった。このことから、提案手法の方が MOTA、IDF1 両者において高い数値を示していることがわかる。

図 2 に提案手法による人物移動軌跡の推定結果を示す。これらの図において、人物を囲う枠の色が同一の場合は、同一人物として推定されたことを表す。図 2(a)では、水色枠の人物が椅子に座っており、人物検出が正しくできている。図 2(b)では、赤色枠の人物が画像の右端から左に向かって歩いている様子が検出されており、椅子に座っていた人物は遮蔽されてしまっているため検出失敗していることが確認できる。このとき、時間的に離れた動線片同士を関連付ける必要がある。図 2(c)では、図 2(a)における水色枠の人物と図 2(b)における赤色枠の人物両方が正しく検出されている。つまり、提案手法によって各人物ごとに正しく同一人物として判定できていることがわかる。

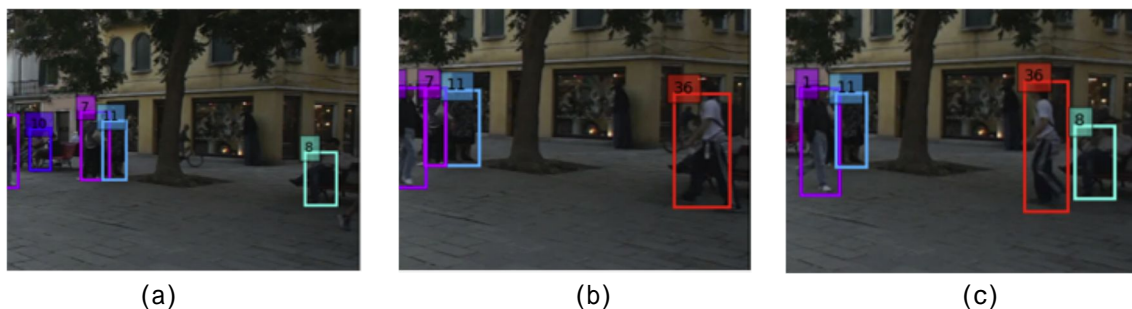


図 2 人物移動軌跡推定結果 (左から時刻 t=1, 10, 20)

(2) 動作特徴を用いた個人認証

階層的瞬き動作認証の性能評価

各被験者について、近距離(カメラからの距離 1m)、遠距離(カメラからの距離 2m)において瞬き動作をそれぞれ 50 回ずつ撮影した。これらを近距離瞬き動作映像、遠距離瞬き動作映像とそれぞれ呼ぶ。被験者の人数は 10 名とした。

本実験では、以下の項目について提案システムの性能評価を行った。

- 遠距離瞬き動作映像の直接識別： 近距離瞬き動作映像を用いて訓練した識別器を用い、遠距離瞬き動作映像を入力した時の識別性能を確認した。
- 開発したシステムによる階層的識別： 近距離瞬き動作映像を用いて、階層的識別器を構築した。遠距離瞬き動作映像をグループ識別に、近距離瞬き動作映像をグループ内個人識別に入力したときの識別性能を確認した。

実験により得られた平均個人識別率を表 2 に示す。階層的識別による結果が、直接識別により得られた結果よりも 16.8%高いことが確認できる。このことから、階層的識別に基づくアプローチに優位性があることがわかる。

表 2 平均個人識別率の結果比較

学習	テスト	グルーピング	平均個人識別率
近距離	遠距離	なし	18.00%
近距離	遠距離	あり	34.80%

しかしながら、平均識別率は 34.80%であり、生体認証において望まれる識別率としては低い。これは、近距離・遠距離で撮影された瞬き動作の特徴が互いに異なるため

あると考えた。具体的には、被験者の注視位置の違いに起因するものと考えた。

そこで、近距離・遠距離における瞬き動作の撮影において被験者の注視位置を同様にするために、撮影中は特定位置に配置した物体を注視するように指示した。このような環境において、再度評価実験を行った。

結果、注視位置を固定しない場合は平均識別率が 64.0%、注視位置を固定した場合は平均識別率が 84.0%であった。このように、注視位置を固定した場合の方が、個人識別率が向上した。以上より、瞬き動作認証では、被験者の注視点位置を考慮することが重要であることがわかった。

顔特徴と読唇処理を用いた階層的個人認証の性能評価

開発した階層的認証手法の有効性を確認するために、顔認証、読唇に基づく認証、階層的認証（本手法）の性能比較を行った。評価尺度として、等価エラー率（EER）を用いた。EER は、数値が小さい値である方が高性能であることを意味する。

実験の結果、各手法による認証により得られた EER はそれぞれ、16.7%、6.7%、3.3%であった。この結果から、開発した階層的認証手法の有効性を確認することができた。

参考文献

1. C. Weihua, C. Lijun, C. Xiaotang, and H. Kaiqi, "An Equalized Global Graphical Model-Based Approach for Multi-Camera Object Tracking", *IEEE TCSVT*, vol. 27, no. 11, pp. 2367-2381, 2017.
2. B. Alina, S. Denman, S. Sridharan, C. Fookes, and P. Lucey, "A Database for Person Re-Identification in Multi-Camera Surveillance Networks", in *Proc. IEEE DICTA*, pp. 1-8, 2012.
3. T. Matsukawa, T. Okabe, E. Suzuki and Y. Sato, "Hierarchical Gaussian Descriptor for Person Re-Identification", in *Proc. IEEE CVPR*, pp. 1363-1372, 2016.
4. M. Andriluka, S. Roth and B. Schiele, "People-Tracking-by-Detection and People-Detection-by-Tracking", in *Proc. IEEE CVPR*, pp. 1-8, 2008.
5. M. Andriluka, S. Roth and B. Schiele, "Monocular 3D Pose Estimation and Tracking by Detection", in *Proc. IEEE CVPR*, pp. 623-630, 2010.
6. H. Shen, L. Huan, C. Huang and W. Xu, "Tracklet Association Tracker: An End-to-End Learning-based Association Approach for Multi-Object Tracking", *arXiv preprint arXiv:1808.01562*, 2018.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 杉村大輔, 中野真理子, 徳永律子, 内田葉子	4. 巻 31
2. 論文標題 多段階識別に基づく瞬き動作認証手法の検討	5. 発行年 2020年
3. 雑誌名 日本工業出版(株), 画像ラボ	6. 最初と最後の頁 32-36
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計4件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 中野真理子, 徳永律子, 内田葉子, 杉村大輔
2. 発表標題 遠距離・近距離における多段階瞬き動作認証
3. 学会等名 電子情報通信学会画像工学研究会
4. 発表年 2019年

1. 発表者名 稲見慎吾, 杉村大輔, 浜本隆之
2. 発表標題 集団・個人の見え特徴の併用によるカメラ間人物照合
3. 学会等名 電子情報通信学会画像工学研究会
4. 発表年 2019年

1. 発表者名 渡邉丈裕, 前田慶博, 杉村大輔, 浜本隆之
2. 発表標題 人物移動軌跡推定のためのグラフニューラルネットワークに基づくリンク予測
3. 学会等名 映像メディア処理シンポジウム
4. 発表年 2021年

1. 発表者名 安永綾花, 岡田祐花, 園櫻子, 杉村大輔
2. 発表標題 顔特徴と発話時の唇の動きに基づく階層的生体認証
3. 学会等名 電子情報通信学会総合大会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	浜本 隆之 (HAMAMOTO Takayuki) (10297624)	東京理科大学・工学部電気工学科・教授 (32660)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------